GATE LYON SAINT-ÉTIENNE
UMR 5824

93, chemin des Mouilles
69130 Ecully - France

Maison de l'Université, Bâtiment B
10, rue Tréfilerie
42023 Saint-Etienne cedex 02 - France

http://www.gate.cnrs.fr
gate@gate.cnrs.fr

WP 2018 – May 2020

Evaluating the impact of public policies on large firms: a synthetic control approach to science industry transfer policies

Corinne Autant-Bernard, Ruben Fotso, Nadine Massard

Abstract:

Large firms dominate R&D investment in most countries and receive the majority of public R&D funding. Due to methodological difficulties, however, evaluation of the effect of government-sponsored R&D programmes mainly focuses on small- and medium-sized enterprises. The scarcity of large firms and their heterogeneity hampers the ability to find proper counterfactuals for very large companies and makes it difficult to use proper inference methods to measure the impact of a specific policy. In order to address these methodological issues, we propose using the synthetic control method, initially developed by Abadie et al. (2010) to evaluate programmes on a regional scale. We apply this method to evaluate the impact of a new French science-industry transfer initiative and compare the results with the random trend model and more standard counterfactual approaches. Based on data covering a long pre-treatment period (1998–2011) and ongoing treatment period (2012–2015), we reveal a convergence between the results obtained with the synthetic control method and the random trend model, and demonstrate that traditional counterfactual evaluation methods are not appropriate for large firms. Moreover, the synthetic control method has the advantage of providing an individual assessment of the policy impact on each firm. In the specific case of the French science-industry transfer initiative, it reveals that the impact on private R&D is highly heterogenous both on RD inputs and cooperation behaviours. Beyond this specific transfer policy, this study suggests that the synthetic control method opens new research perspectives in policy impact evaluation at the firm level.

Keywords:

impact evaluation, R&D policy, large firms, synthetic control method, Technological Research Institutes (TRIs)

JEL codes:
C23, D22, O36, O38

Working Paper

Evaluating the impact of public policies on large firms: a synthetic control approach to science-industry transfer policies

Corinne Autant-Bernard,[1] Ruben Fotso[2] and Nadine Massard[3]

May 2020

Abstract:

Large firms dominate R&D investment in most countries and receive the majority of public R&D funding. Due to methodological difficulties, however, evaluation of the effect of government-sponsored R&D programmes mainly focuses on small- and medium-sized enterprises. The scarcity of large firms and their heterogeneity hampers the ability to find proper counterfactuals for very large companies and makes it difficult to use proper inference methods to measure the impact of a specific policy. In order to address these methodological issues, we propose using the synthetic control method, initially developed by Abadie *et al.* (2010) to evaluate programmes on a regional scale. We apply this method to evaluate the impact of a new French science-industry transfer initiative and compare the results with the random trend model and more standard counterfactual approaches. Based on data covering a long pre-treatment period (1998–2011) and ongoing treatment period (2012–2015), we reveal a convergence between the results obtained with the synthetic control method and the random trend model, and demonstrate that traditional counterfactual evaluation methods are not appropriate for large firms. Moreover, the synthetic control method has the advantage of providing an individual assessment of the policy impact on each firm. In the specific case of the French science-industry transfer initiative, it reveals that the impact on private R&D is highly heterogenous both on RD inputs and cooperation behaviours. Beyond this specific transfer policy, this study suggests that the synthetic control method opens new research perspectives in policy impact evaluation at the firm level.

Keywords: impact evaluation, R&D policy, large firms, synthetic control method, Technological Research Institutes (TRIs).

JEL codes : C23, D22, O36, O38

[1] Univ Lyon, UJM Saint-Etienne, GATE UMR 5824, F-42023 Saint Etienne, France.

[2] Univ Lyon, UJM Saint-Etienne, GATE UMR 5824, F-42023 Saint Etienne, France.

[3] Université Grenoble Alpes, CNRS, INRA, Grenoble INP, GAEL, 38000 Grenoble, France.

I. Introduction

Due to their weight in the economy, large firms are often the first to benefit from public funding. A comprehensive database of corporate subsidies in the US indicated that a relatively small number of large companies receive most federal grants and are allocated most tax credits.[4] Of the $68 billion awarded between 2000 and 2015, two-thirds went to fewer than 600 large companies and twenty-one of them netted more than $500 million in federal grants and subsidies. Such a tendency is observed in almost all public policy areas. However, this appears puzzling when considering innovation policy, since the factors that hamper innovation are often more striking in small firms.

Large firms dominate R&D investment in most countries. Only 250 companies, according to the OECD Science, Technology and Industry Scoreboard (2015), perform more than 60% of global R&D. Moreover, since most public funding is proportional to private R&D spending, a small number of large corporate players receive the largest share of public funding. For instance, in 2016, large French companies[5] received around 73% of direct public R&D support whereas SMEs and mid-cap companies respectively received 20% and 7% (source: MHER, 2018[6]). Despite the importance of these large actors, literature on the quantitative microeconomic evaluation of the effect of government-sponsored R&D programmes mainly focuses on small- and medium-sized enterprises (Dimos and Pugh, 2016). Many arguments are put forward to justify the focus of evaluation studies on SMEs, among them it is often argued that smaller firms face greater obstacles to innovate while being important job generators. Only very few studies include both small and large firms (Criscuolo *et al.* 2012, Bondonio *et al.* 2015, Dujardin *et al.* 2015, Ben Hassine and Mathieu, 2017, Greco et al., 2017, Vanino *et al*., 2019). They provide either no specific assessment or inconclusive results on the impact of innovation policies on large firms. Although there are some theoretical explanation for the limited impact of R&D policies on large firms (the lower share of aid in the R&D investments of big companies or the higher risk of opportunistic behaviour), it is often methodological problems that prevent the production of significant and robust statistical results of impact evaluation.

Indeed, due to the scarcity and heterogeneity of large firms, it is difficult or impossible to find proper counterfactuals for them and thus to predict what would have happened without public intervention, as well as to make proper inferences on the impact of the programme. This paper aims to tackle these methodological issues.

We argue that the synthetic control method (SCM) developed by Abadie and Gardeazabal (2003), and Abadie *et al.* (2010) to assess policy impacts at the regional or country level may constitute a relevant

[4] Uncle Sam's Favorite Corporations: Identifying the Large Companies that Dominate Federal Subsidies, Philip Mattera and Kasia Tarczynska, March 2015, Good Jobs First, www.goodjobsfirst.org.

[5] Firms with more than 5,000 employees or more than 1.5 billion euros in turnover and more than 2 billion euros in total assets. They are 274 such firms in France.

[6] French Ministry of Higher Education and Research.

tool for evaluating the effectiveness of public policy on large firms. As stated by Athey and Imbens (2017), the SCM is arguably the most important innovation in the policy evaluation literature in the past 15 years. This method builds on an estimation of difference-in-differences, but systematically uses more attractive comparisons. This approach offers several advantages. Firstly, it moves away from using a single control unit or a simple average of control units, and instead uses a weighted average of the set of controls, allowing us to build a counterfactual situation for very large firms. Secondly, the SCM provides individual impact evaluation for each firm, rather than computing an average treatment effect, which allows taking full account of inter-individual heterogeneity. Thirdly, this quantitative approach could be combined with more qualitative information in order to understand the sources of heterogeneity and therefore to improve the design of public policies.

The paper first discusses the methodological difficulties faced when it comes to evaluating the impact of public intervention on large firms. The solutions offered by the SCM are then analysed and illustrated through empirical application to a science-industry transfer policy. Such programmes aim not only at fostering private R&D investment but also at promoting collaborative behaviours. Large enterprises are often at the heart of science-industry transfer programmes and public funding is mainly justified as helping to remove obstacles due to the collective organisation of R&D and the transformation of research into open and valuable innovations. The French Technological Research Institutes (TRIs) programme is an iconic example of such policy and offers a relevant field of application to demonstrate the advantages of SCM by assessing the impact of this programme on the R&D behaviour of large firms. Derived from one of the largest French investment programmes implemented in 2010, TRIs are interdisciplinary research institutes which bring together private actors (SMEs, large companies, etc.) and public actors (public research organisations, universities, etc.) around a strategic research programme, encouraging them to work together through technological platforms in order to accelerate the transfer of knowledge to industry. Large companies play a key part in this process, as they invest themselves in the platforms that small firms can then use. They provide physical and intangible assets and human resources that interact with scientists in the public sector. This strong involvement in the technological platform is expected to encourage knowledge transfer between public research labs and large firms and therefore to boost firms' investment in R&D.

Our evaluation relies on data covering a long period of pre-treatment (1998–2011) and ongoing treatment (2012–2015). Comparisons between the SCM, the random trend model and traditional counterfactual evaluation methods, especially the double-difference method, and the fixed effects model, teach us that the latter two methods lead to biased results, probably because they fail to take into account certain important characteristics of large firms, in particular the long-term effect of their strategies (mergers, acquisitions, etc.) and the trend of the input indicators of these companies. The random trend model offers an interesting alternative although continues to be limited for statistical inference on such a small sample. The SCM has the advantage of providing an individual assessment of the policy impact on each firm.

In the specific case of our science-industry transfer programme, the SCM reveals no systematic impact. A leverage effect was found on the R&D inputs of one firm, but a negative effect was observed on another, suggesting a windfall strategy of this participating firm. A significant positive impact of the programme on R&D cooperative behaviour was also only revealed for one company, as this firm shows a higher tendency to develop external R&D compared to its synthetic control. Such heterogeneous individual results open the way toward a more comprehensive qualitative interpretation of the effectiveness of the programme by considering the role of firms' characteristics and strategies as well as the way this new platform-based science-industry initiative is implemented.

This paper therefore contributes to the literature in three ways. From a methodological perspective, it suggests that the SCM could be a useful approach to evaluating the policies from which large firm benefit. It could therefore, open new avenues for research into policy impact evaluation at the firm level. From a policy perspective, we apply this method to a new science-industry transfer policy. Based on technological platforms, this policy provides new incentives and a favourable context for firms to access and exchange knowledge with public sector labs. Recently implemented in France, this policy has not yet been evaluated. Finally, from an interpretative perspective, it appears relevant to combine the quantitative SCM with more qualitative information in order to understand the heterogenous effects of the policy. Although in our case, firm-specific information is not exploitable due to confidentiality issues, some specificities of the technological platforms may explain their varying ability to increase large firm R&D investments.

The rest of the paper is organised as follows. *Section II* reviews the most significant methodological difficulties that hamper the quantitative impact evaluation of R&D public policies for large companies. *Section III* then presents how the SCM could solve these difficulties. *Section IV* goes on to describe the French TRI programme and our identification strategy, and *Section V* presents how the SCM is applied in practice, before concluding with some reflections on the advantages and disadvantages of using the SCM compared to the other main quasi-experimental methods of impact evaluation. *Section VI* concludes with both the policy evaluation results and methodological issues, calling for a systematic evaluation of the impact of public policies on large firms.

II. Why is quantitative impact evaluation difficult when it comes to large firms?

When randomised experiments are unfeasible, which is the case most of the time when it comes to evaluating economic policies, quasi-experimental designs can be exploited to evaluate causal treatment effects (Imbens and Rubin, 2015). The principle of the quasi-experimental evaluation method consists in finding individuals who are not affected by the measure, but who are as comparable as possible in terms of their initial characteristics to those who have benefited from it (known as 'treated' individuals). These non-recipients are referred to as the 'control group'. They make it possible to estimate what would have been observed for the beneficiaries of the policy had it not been implemented, and therefore make it possible to estimate the effect of the treatment, reducing bias due to confounding factors. In other

words, they make it possible to reconstruct what is known as the 'counterfactual' situation. Measuring the impact of a public policy then compares this virtual scenario without the policy with what can be observed when the policy has been put in place. This comparison can be made using a single difference between the two groups (beneficiaries and non-beneficiaries) or using double-differences (variation observed for beneficiaries compared to variation observed for non-beneficiaries).

Such traditional programme evaluation methods rely on a number of assumptions necessary to identify a causal effect (see Angrist and Pischke, 2008). However, these assumptions are rarely verified for large companies. This is due to two characteristics of very large firms: their small number and their great heterogeneity.

II.1. Difficulty building the counterfactual

The main risk of quasi-experimental methods is the lack of similarity between the beneficiary group (treated) and the non-beneficiary group (controls). Before the intervention, the control and treatment groups should be as similar as possible in order to estimate the effect of the treatment and reduce bias due to confounding factors. While there are many methods for constructing this counterfactual (most of them relying on the propensity score[7]), these are problematic when considering very large companies. A low number of observations reduces the possibility of matching. It is unlikely that when only a few dozen observations are available that a valid counterfactual will be found for each of the treated individuals. In this regard, it is also important to remember that the robustness of the counterfactual method comes from the fact that several controls are available for the same treated observation.

This difficulty is compounded by the heterogeneous nature of large companies and by the fact that, very often, only one or a few very large firms dominate a given sector of activity (BEIS, 2017). It is, therefore, very difficult to find a valid counterfactual for these companies. In addition, in many cases, most large firms benefit from the policy, reducing the possibility of finding non-treated firms.

For all these reasons, traditional methods of constructing a counterfactual are not suitable for large companies. The use of panel data overcomes some of these limitations. By introducing individual fixed effects, it is possible to take into account certain specificities of each company. However, these individual effects, which may be due to managers' strategies within a specific corporate culture, are likely to vary over time. The introduction of fixed effects is therefore not sufficient.

II.2. Difficulty forecasting what would have happened without the policy

The difference-in-differences method (DiD), which consists of differentiating between the situations after and before the policy in each of the groups (beneficiaries and non-beneficiaries) and identifying the difference, is only robust if it is assumed that there is a common trend between the two groups. Under this assumption, the relevant indicators (e.g. the company's economic performance) follow the same

[7] See Stuart (2010) for a review.

trajectory for the policy beneficiary group and the control group. However, this hypothesis is quite strong in the case of large companies due to their high heterogeneity. Large companies are often subject to previous individual shocks (acquisitions, disposals, etc.) that will not be without consequences on their current performance. Firms' strategic behaviours can therefore make it difficult to identify the specific effects of the policy. Moreover, large firms may benefit from several programmes at the same time, which may prevent identification of the specific effect of each.

II.3. Difficulty making inferences

The standard evaluation methods (DiD or instrumental variables) are based on the Gaussian model. Due to the small number of observations when analysing very large firms, it is difficult to consider that the law of the sample is Gaussian. Therefore, a linear regression model cannot be used to assess the impact of a programme.

In addition, given the very strong heterogeneity of large companies, it is likely that the impacts of the programmes will be specific to each firm. Traditional methods of programme impact assessment could enable an explicit consideration of heterogeneous effects. However, these methods require the identification of a variable at the origin of this heterogeneity and they do not make it possible to evaluate differentiated effects for each individual. Once again, the use of panel data may offer interesting improvements. Fixed effects can capture the individual specificities of companies. However, it is important to go beyond this in order to be able to measure not an average effect but an individual effect of the policy.

For all these reasons, traditional programme evaluation methods are not suitable for large companies. It is, therefore, important to find appropriate methods. The synthetic control method, initially developed by Abadie *et al.* (2010) to evaluate programmes on a regional scale, may offer interesting solutions.

III. How the synthetic control method (SCM) could solve the difficulties

III.1. Building counterfactual situations with the SCM

Unlike counterfactual evaluation methods, which aim to assess the average effect of the treatment, the SCM has the particularity of evaluating the individual effect of the treatment. The main challenge of this method is to construct the counterfactual of a single treated unit. According to the SCM (Abadie *et al.*, 2003), the counterfactual of a single treated unit is estimated by the weighted linear combination of untreated companies that best reproduces the characteristics of the treated unit over time, prior to the treatment. The principle of counterfactual construction is to find the weight (or proportion) that minimises the distance between the values of the pre-treatment variables of the treated unit and the values of the same variables of the other untreated units. In other words, untreated units must be selected so that their pre-treatment characteristics correspond as closely as possible to those of treated unit. To

do this, all the untreated units are considered to belong to a pool of potential counterfactual units that Abadie and Gardeazabal (2003) called the "donor pool", and each of the potential counterfactual units is supposed to contribute in a certain proportion (weight) to the reconstruction of the values of the pre-treatment variables for the treated unit. According to Abadie *et al.* (2003) and Abadie *et al.* (2010), the optimal weights should be chosen by minimising the mean squared prediction error of the results before treatment.

To better understand this approach, following Abadie *et al.* (2015), let us suppose a sample of J+1 units that are observed, where j = 1 is the unit that is exposed to treatment (intervention, programme, etc.), so that the remaining J units (from j = 2 to j = J+1) are considered as potential counterfactual units, i.e., the units constituting the donor pool.

Now, suppose that the sample is a balanced panel, i.e., a set of longitudinal data where all units are observed at the same time period, t = 1 ... T. The study period is divided into two periods: a pre-treatment period that includes a number of periods T_0 and a post-treatment period that includes a number of periods T_1 so that $T = T_0 + T_1$. In order to simplify the formalisation, suppose that the treated unit "1" is exposed to the intervention during the periods $T_0 + 1$, ..., T, and that the intervention has no effect during the pre-treatment period 1, ... , T_0.

Let us now consider a weight vector of dimension (J x 1), $W = (w_2, \ldots w_{J+1})'$ such as $0 \leq w_j \leq 1$ and $w_2 + \ldots + w_{J+1} = 1$. Each particular value of the vector W represents the weight of unit j in the synthetic control (Abadie *et al.*, 2010). Therefore, the choice of a subset of valid control units depends on the values of the vector W. Let V be a diagonal matrix with non-negative values. The values of the diagonal elements of V reflect the relative importance of different predictors (pre-treatment characteristics) of the treated unit. Suppose that the weight vector W* represents the set of optimal values that minimise the distance between the values of the pre-treatment variables of the treated unit and those of the control units. Let us designate the variable Y_{1T} as a vector of a dimension (T x 1) containing the values of the outcome variable of the treated unit during T periods and Y_{JT}, a matrix of dimension (T x J) containing the values of the same outcome variables for the J control units. Let us designate by X_{1T}, a dimension vector (KT x 1) of the pre-treatment values of the predictive variables associated with the outcome variable Y_{1T} and $X_{JT,}$ a dimension matrix (KT x J) which contains the values of the same variables for the J potential control units. Note that the predictive variables must not be affected by the intervention.

The determination of the optimal weights W* is carried out in two phases of the optimisation process (Kaul *et al.* 2018). The first phase called "inner optimisation" consists of searching a combination of untreated units in the donor pool so that the difference between the values of predictive variables of the treated unit and the control units is as small as possible. This difference is measured by the following metric distance:

$$\left|\left|X_{1T_0} - X_{JT_0}W\right|\right| V = \sqrt{\left(X_{1T_0} - X_{JT_0}W\right)' V \left(X_{1T_0} - X_{JT_0}W\right)} \quad (1)$$

Thus, for any given prediction weight V, internal optimisation consists in finding the non-negative weights W of the untreated units summing up to the unit ($w_2 + ... + w_{J+1} = 1$) such that:

$$\sqrt{\left(X_{1T_0} - X_{JT_0}W\right)' V \left(X_{1T_0} - X_{JT_0}W\right)} \rightarrow min. \quad (2)$$

The solution to this problem is given by $W^*(V)$.

Once the optimal weights W^* are determined as a function of V, the second phase of the optimisation process called "outer optimisation" consists in finding the optimal weights of the predictive variables. Following a data-driven approach, Abadie and Gardeazabal (2003) and Abadie *et al.* (2010) propose choosing V from all defined positive and diagonal matrices such that the mean squared prediction error of the outcome variable is minimised during the pre-treatment periods. The mean squared prediction error (MSPE) is given by:

$$MSPE(Y) = \frac{1}{T_0}\sum_{t=1}^{T_0}\left[\left(Y_{1t} - \sum_{j=1}^{J} w_j^*(V) Y_{jt}\right)^2\right] \quad (3)$$

Thus, the results of the "outer optimisation" problem[8] are given by:

$$\left(X_{1T_0} - X_{JT_0}W^*(V)\right)'\left(X_{1T_0} - X_{JT_0}W^*(V)\right) \rightarrow min \quad (4)$$

V is thus selected to weight covariates in accordance with their predictive power on the outcome.
After determining the optimal weights, we sought to estimate the counterfactual trajectory of the treated unit, i.e., the trajectory of the outcome variable that the treated unit would have known in the absence of treatment.

III.2. Using the SCM to forecast what would have happened without the policy

The SCM is based on the principle that if the synthetic control provides a good approximation of the result for the treated unit over a long pre-treatment period, then any subsequent difference between the treated unit and the synthetic control could be attributed to the effect of treatment on the outcome (identification assumption).
The counterfactual trajectory of the outcome variable is estimated by the weighted linear combination of the same outcome variables of the untreated units. The estimated counterfactual of the outcome variable of the treated unit is thus given by:

$$Y_{1t}^* = \sum_{j=2}^{J+1} Y_{jt} w_j^*, \quad (5)$$

[8] In practice, the SYNTH package of the R software automatically provides the optimal vector weights corresponding to the solutions of the two optimisation phases.

where *W** is the vector of optimal weights which minimise the distance between the values of the pre-treatment variables and those of the same pre-treatment variables of untreated units.

To assess the impact of the intervention on the treated unit, we simply compare the results of treated unit after treatment with those that the unit would have achieved in the absence of treatment.

This assumes, as is standard in counterfactual approaches, that the outcome variable Y_{1t} of the treated unit can take two possible values: Y_{1t}^1, the value observed after treatment and Y_{1t}^0, the value that it would have obtained in the absence of treatment.

$$Y_{1t} = \begin{cases} Y_{1t}^1 \\ Y_{1t}^0 \end{cases} \quad \text{For t = 1, 2, … } T_0, T_0+1, T_0+2, \ldots T \tag{6}$$

The objective is therefore to estimate the annual effects of treatment: $\alpha_{1T_0+1}, \alpha_{1T_0+2}, \ldots \alpha_{1T}$

Where $\alpha_{1t} = Y_{1t}^1 - Y_{1t}^0$ (7)

By definition, Y_{1t}^0 is not observable, in the sense that it is the value that the variable Y_{1t} would have recorded if the treated unit had not been treated. In other words, it is the counterfactual situation of the treated unit.

Suppose now a general model for the potential results of all units such that the outcome variable of unit *i* at time *t* is given by:

$$Y_{it} = Y_{it}^0 - \alpha_{it} D_{it} \tag{8}$$

With *i* = 1, … J+1 and D_{it} is the indicator of treatment that takes the value 1 if the unit *i* is treated and *t* > T_0 and 0 otherwise. Abadie *et al.* (2010) formulate the variable Y_{it}^0 by using the following linear model:

$Y_{it}^0 = \delta_t + \theta_t X_i + \gamma_t \mu_i + \varepsilon_{it}$; for all *i* = 1, … J+1 and for all *t* = 1, … T. (9)

where δ_t is a vector of common time-specific effects, with constant individual effects between units, θ_t a vector of unknown time-specific parameters, X_i the observed predictive variables that are relevant to the outcome variable. μ_i an unobservable effect specific to the unit *i*, γ_t an unknown common factor. ε_{it} a not observed transitory shock at level *i* with a zero average for any conditional i (δ_t, X_i , μ_i).

By using the linear factor model just described in (9), Abadie *et al.* (2010) show that if the number of pre-intervention periods is large relative to the scale of transitory shocks, we can choose W* such that:

$$\sum_{j=2}^{J+1} w_j^* Y_{jT_0} = Y_{1T_0} \quad and \quad \sum_{j=2}^{J+1} w_j^* X_j = X_1 \tag{10}$$

$$\text{Then, } \hat{\alpha} = Y_{1T_1} - \sum_{j=2}^{J+1} w_j^* Y_{jT_1} \tag{11}$$

is an unbiased estimator of α for the T_0+1, ... T periods, i.e., the impact of the intervention. Thus, the synthetic control estimator of the treatment effect is given by comparing the post-intervention results between the treated unit, which was exposed to the intervention, and the synthetic control, which was not exposed to the intervention. In other words, for each post-intervention period *t* (where $t \geq T_0$), the synthetic control estimator of the treatment effect is given by comparing the result of the treated unit with the result of the synthetic control in that period.

The estimator of annual effects on the treated unit "1" at time *t* is given by:

$$\hat{\alpha}_{1t} = Y_{1t}^{1} - \sum_{j=2}^{J+1} w_j^* Y_{jt}, \text{ for } t = T_0+1, \ldots T \quad (12)$$

It is also possible to estimate the average annual effect of treatment on the treated unit by:

$$\hat{\alpha}_1 = \frac{1}{T-T_0} \sum_{t=T_0+1}^{T} \left[Y_{1t} - \sum_{j=2}^{J+1} w_j^* Y_{jt} \right] \quad (13)$$

Note that a note of caution should, however, be sounded here. Implementation of the SCM often faces two main problems: interpolation bias and over-adjustment.[9] 'Interpolation' bias refers to the distance between the function that generates the values of outcome variable for the treated unit and the one approximated using the SCM. These biases become problematic when they are severe, which they can be in some cases, particularly if the donor pool contains units with very different characteristics from those in the treated unit or if the relationship between the outcome variable and the explanatory variables in X_{1t} and X_{jt} is highly non-linear (Abadie *et al.*, 2010).[10]

III.3. Using the SCM to draw inferences

It is difficult, or even impossible, to apply traditional statistical inference approaches to comparative case studies due to the small sample size, lack of randomisation and the fact that probability sampling is not used to select sampling units (Abadie *et al.*, 2015). Therefore, the SCM proposes other inference methods. To assess the significance of the estimates, the research question is whether the results obtained could be entirely attributable to chance. To answer this question, it is recommended conducting placebo studies. Two main types of placebo studies are identified: "in-time placebo" and "in-unit placebo" (Abadie *et al*., 2010). In the first, the exercise consists in replacing the date of the intervention with a date on which the intervention did not take place (Heckman & Hotz 1989). If the SCM provides a proper counterfactual, a null effect should be obtained at this date, otherwise this method did not correctly predict the trajectory of the results. Abadie *et al.* (2015) and Saia (2017) have applied this test. The second placebo test relies on the same principle by using a fake treated unit selected randomly from the donor pool. It has been used in several studies including Abadie and Gardeazabal (2003), DiNardo and Pischke (1997), Angrist and Krueger (1999), Auld and Grootendorst (2004), Abadie *et al.* (2015), and Saia (2017). Similarly, Abadie *et al.* (2010) propose implementing an extensive version of the placebo test in relation to the unit. Their idea of the placebo test is similar to the classic permutation inference framework, where the distribution of a test statistic is calculated under random permutations of the sample unit assignments to the intervention and non-intervention groups. As in the permutation tests, they apply the SCM to each untreated unit in the donor pool. This makes it possible to assess whether the effect estimated by the synthetic control for the unit affected by the intervention is significant,

[9] For more details, see Abadie *et al.* (2015).

compared to the effect estimated for a randomly selected unit. Abadie *et al.* (2010) implemented this last version which we also rely on in the empirical part.

IV. Empirical analysis and identification strategy

IV.1. The Technological Research Institute policy and its expected impacts

In the aftermath of the 2008 crisis, which greatly affected France and its partners, the French government wanted to go beyond economic responses and carry out structural action by launching a vast investment programme known as the "Programme d'Investissement d'Avenir"[11]. With a current budget of 57 billion euros, this programme focuses on several major areas including higher education, training and research; industrial sectors and SMEs; the digital economy; sustainable development; etc. More specifically, this programme defines several actions, including the creation of the Technological Research Institutes (TRIs) in 2012. Focussing on relationships between science and industry, these TRIs are justified by the particular context that is commonly known as the "European paradox". This paradox refers to the lack of efficiency in innovation despite relatively high levels of public and private research. Indeed, various OECD reports reviewing national innovation policies in France have revealed gaps in the exchange of knowledge between science and industry. These TRIs emerged as a response to this particular situation. The cost to the taxpayer of the eight TRIs created in France over eight years (2012-2020) was 2 billion euros, reflecting the expectation that this initiative will bring substantial rewards for French society that are important to assess.

One of the specificities of the TRIs compared to existing devices lies in their ability to make heterogeneous actors work together (companies and start-ups on the one hand and universities and public research organisations on the other) through multilateral (as opposed to bilateral) cooperation, co-located, based on technological platforms and high-level research teams. Fifty percent of TRI activities are financed by public sector funding and 50% comes from private actors. The private actors are mainly very large firms that support TRIs with cash, qualified human resources, and/or machines and equipment. Such public-private co-investments should enable large companies to acquire higher levels of technology. Risk sharing is also likely to encourage breakthrough innovations, enabling strategic markets to be conquered. In addition, training students through these platforms should make it possible to offer companies a high-level workforce that is better adapted to companies' needs. These long-term partnerships between large companies and higher education and research institutions are therefore expected to bring economic benefits to large companies in terms of R&D, innovation and, in the long term, market share and employment. Due to the lack of temporal hindsight, the impacts that we can expect to observe at this time are mainly those regarding R&D investments. At the R&D stage, TRIs face significant challenges, as they have been clearly assigned the objective of creating the conditions

[11] This literally means "Investment in the Future Programme".

for closer collaboration between private and public research. Two types of leverage effects are therefore to be expected. First, a classical input additionality effect, which consists in comparing policy expenditures with the additional amount of R&D spent by private firms. When policy expenditure is more than compensated for by the additional amount of business R&D spending, this describes what is known as a "crowding-in" or "leverage" effect, indicating a complementarity between public and private funds. Second, a behavioural additionality effect is also expected. As defined by Clarysse *et al.*, 2009 behavioural additionality is related to changes in processes that take place within the firms that benefit from public policies. More precisely concerning TRIs, beneficiary firms are expected to change their R&D collaboration behaviour by developing interactions with other actors in the TRI, particularly public ones. Large firms may invest in R&D activities while using the platform. They may however, also choose to outsource part of their research through this platform to public research labs. More broadly, the new research projects established as a result of the platform might increase their needs for complementary outsourced R&D. It is worth noting that these two objectives are not trivial for TRIs, especially when this concerns large firms. Indeed, in the extant literature, the very few studies that take into account the size of companies often show no significant impact of R&D policies on large companies. Authors explain this absence of significant additional effects either by the low level of public support compared to the total amount of the firm's R&D investment, or by the greater risk of opportunistic behaviour by these companies (Jugend et al., 2020). Moreover, estimating how distinct firms' characteristics influence the realisation of behavioural additionalities (Wansenböck *et al.*, 2013) shows that larger firms which are more R&D intensive are also less likely to substantiate behavioural additionalities (measured by project and cooperation behaviour in particular). In the same vein, Greco et al. (2017) show how difficult it is to define efficient public subsidies to foster open-innovation dynamics, especially for large firms. However, these conclusions are not very informative for the public authorities, since it is difficult to determine whether the lack of significant impact is due to the behaviour of the firms or merely to the statistical difficulties of measuring the impact for this type of firm.

Our objective in this study is to use the proper method to provide more refined results on the influence of the TRIs on the behaviour of large recipient firms. Specifically, we consider three outcome variables (*total R&D expenditure net of public funds, R&D self-financing, and external R&D expenditure*). The first two are indicators of input additionality. Total R&D expenditure net of public funds is an interesting indicator to capture the R&D efforts induced by participation in TRIs, while neutralising the monetary support granted to large companies. Self-financing of R&D is an essential variable that will make it possible to verify whether the estimated impact is not a windfall effect for the company but rather a leverage effect. The third indicator of external R&D expenditure, is an indicator of behavioural additionality as it may be used as a proxy for collaborative practices in R&D.

In order to assess whether a leverage effect on R&D occurs, we put special emphasis on two TRIs, namely Nanoelec and Bioaster, located in the Rhone-Alpes region of France. Two large firms are involved in these two TRIs: STmicroelectronics and Schneider electrics in Nanoelec, and Sanofi Pasteur and Biomérieux in Bioaster. As part of a regional project, interviews were conducted with these two

platforms and we were given specific access to information about them. As detailed below, this made it possible to precisely identify the four large companies involved in these two TRIs as private founding members.

IV.2. Identification strategy

Our ability to identify whether large firms would have undertaken the same R&D project in the absence of the TRIs faces traditional problems of sample selection and endogeneity in the evaluation of innovation policies (Afcha and Garcia-Quevedo, 2016). The first of these problems, sample selection, arises because it is only possible to observe the performance of firms that are participating in the policy. The second problem is that the variables used to measure the effect of public intervention (e.g. private intervention in R&D) could be endogenously determined if we assume that firms making a greater effort in R&D are more likely to be part of a TRI.

Our identification strategy relies on a comparative interrupted time series design (CITS). We use annual panel data at the company level over the period 1998–2015. TRIs were effectively created in 2012, which gives us 14 years of pre-treatment data (1998–2011) and four years of ongoing treatment data (2012–2015). The duration of our pre-treatment period is relatively similar to that identified in the literature, particularly in the studies of Abadie *et al.* (2010) and Saia (2017), which use periods of 14 and 19 years, respectively. In a later study, Abadie *et al.* (2015) covers an even longer period, with 30 years of pre-treatment. We chose not to extend our period of observation too far in order to avoid a strong disequilibrium between the pre- and post-treatment durations. Our sampling period began in 1998 due to the unavailability of data for the majority of companies before this year and ended in 2015 since this is the last year for which all the databases are available. The sources of information that we used to build our dataset come from five databases: the French annual survey of firms,[12] a French employment database,[13] the R&D survey,[14] the R&D tax credit database,[15] and internal data from the two TRIs in the Rhône-Alpes region.

As stated above, "treatment" is defined as the fact of being a private founding member of the TRI. Thus, the treated units are the four large companies that are private founding members of TRIs. In order to verify that treated companies are effectively involved in the TRIs, for each establishment of these large

[12] This database, known as Ficus&Fare, is produced by the French Ministry of Finance (DGFiP) and INSEE (National Statistical Institute).

[13] Known as DADS, this dataset contains information on employees' positions and individual characteristics and some information on the employer. It is produced by INSEE.

[14] This survey, produced by the French Ministry of National Education, Higher Education and Research, records information on human and financial resources dedicated to R&D.

[15] The GECIR database is produced by the French Ministry of Finance provides information about research tax credits. Data are only available from 2008 to 2014.

companies we identified its location, its total workforce and its proportion of managers. For confidentiality reasons, we cannot publish the detailed results, but they indicate that, on average, 82.23% of the total workforce and 83.14% of the proportion of managers are located in the treated area (the Auvergne-Rhône-Alpes Region). This therefore implies that a high proportion of the large companies are involved in the TRI policy. A well-known aspect for causal effect identification is that treatment and comparison subjects prior to the analysis would not differ in levels and slopes. Within the framework of the SCM, and in order to avoid problems of over-adjustment and interpolation bias, it is ideally recommended to choose untreated companies whose pre-treatment values of the predictive variables and outcome variable are similar to those of the treated company over the entire pre-treatment period. The untreated units that are affected by an intervention of a similar nature or that are likely to have experienced large idiosyncratic shocks to the outcome variables during the study period, should also be excluded, if such shocks would not have affected the treated unit in the absence of the treatment (Abadie *et al.*, 2015). Costa Dias (2014) recommends combining the matching and SCM. For Costa Dias, this combination (matching before applying a SCM) may help ensure the comparability of allowable controls. Matching would be done on the pre-treatment characteristics and the outcome variable during the pre-treatment period. However, it may not be possible to match several characteristics so closely. Therefore, we selected only those firms that had a similar level and growth in the preintervention activities as firms participating in the TRIs.

With this aim, we moderated the selection criteria by introducing filters to build our donor pool sample. The first filter focused on selection variables, the second filter was based on selection margins, and the third was based on years of pre-treatment. For the first, we considered two selection variables, turnover and total R&D expenditure. The choice of these two variables is justified by the need to select companies of a similar size that carry out a volume of R&D activity close to that of the treated companies. Firms investing in TRIs are indeed likely to have higher R&D spending than other firms. More widely, R&D is clearly acknowledged as the main selection variable in the literature on innovation policy evaluation (Boeing, 2016). Moreover, the size of a firm acts as a confounding characteristic that affect both R&D spending and TRI participation. Thus, a firm's participation in a TRI becomes endogenous to the firm's own R&D efforts. In other words, even in the hypothetical absence of TRI participation, the R&D expenditure of participants is likely to be higher than those of non-participants, leading to an overestimation of the actual effect of the policy if size is ignored.

In order to ensure some homogeneity between treated and un-treated firms with regard to these two variables, we set a selection margin equal to plus or minus 33% of the values of these selection variables. Regarding the years of selection, we chose to consider three years over the entire pre-treatment period, corresponding to the beginning, middle and end of this period, which results in the years 1999, 2005 and 2011. Concretely, it is a matter of selecting all the untreated companies with a turnover and total

R&D expenditure for the years 1999, 2004 and 2010 which are similar to those of the turnover and total R&D expenditure of the treated companies[16].

By applying these filters, we were able to constitute a sample of 34 potential untreated companies observed over the 1998–2015 period. In order to be sure that none of the 34 control companies benefited from the actions of the other six French TRIs, a thorough search was conducted on each of these 34 companies to verify that they did not benefit from the actions of the six other TRIs, nor that they were subsidiaries of beneficiary companies. This action led to the removal of 12 companies, which resulted in our sample being 22 control firms. By adding the four treated companies to these 22 control companies, we obtained a balanced panel dataset of 26 companies observed from 1998 to 2015, which represents a set of 468 observations.

In order to verify whether the non-treated companies can constitute an appropriate control group for our treated companies, we carried out a comparative analysis between the treated and the firms of the donor pool sample in terms of pre-treatment characteristics and outcome variables.
As shown in the Table 1 and Figures 2 to 4 in appendix A, in spite of these filters, our treated and un-treated groups still differ greatly in their outcome variables in both scale and trend. Several conclusions can be drawn from these observations. The first confirms that the control group is not a proper counterfactual group if we consider the whole set of donor pool companies. Therefore, it is necessary to determine, as suggested by the SCM, the weighted linear combination of untreated companies that would constitute a valid counterfactual group. The last step of our identification strategy thus lies in building a counterfactual situation using the SCM. This is detailed in the next sub-section. The second conclusion is that difference-in-differences method is not suitable for assessing the effect of TRIs on large companies, since the fundamental assumption of common trends, which implicitly stipulates that the treated and untreated groups must have similar trajectories during the pre-treatment period, is biased. The third conclusion concerns the specification of our model, which should make it possible to take into account both common annual shocks and individual annual shocks.

V. Impact analysis of the TRI programme on large firms

Relying on a comparative interrupted time series design (CITS), our objective is to identify whether large firms would have undertaken the same R&D project in the absence of the TRIs. This is achieved by comparing the results of each treated unit with those that the unit would have faced in the absence of the policy. The first challenge of the synthetic control method is therefore to be able to estimate the counterfactual situation.

[16] When data are missing, variables are observed in 2000, 2005 and 2011.

V.1. Construction of the synthetic controls

According to the SCM, the synthetic control of each company must be constructed by the convex combination of the large companies in the donor pool that most closely resemble this company in terms of pre-TRI values of predictive variables and outcome variables.

To this end, we identified a set of ten predictive variables, which may explain both TRI participation and R&D outcome. It is essential that the choice of pre-treatment characteristics should include variables that have the predictive power of the trajectory of the treated unit, but not those that anticipate the effects of the intervention (Abadie *et al.*, 2010). In addition, predictive variables must not be affected by the treatment. Along with R&D expenditure and turnover already discussed above, we consider three groups of potential selection variables. The first group accounts for firm size, considering indicators in addition to turnover including employment, equity and number of establishments. A second group reflects the technological orientation of the firm, focusing on indicators in addition to R&D expenditure. This includes R&D capital, R&D funding, R&D tax credit and proportion of high skilled labour. Finally, the international orientation of the firm is also accounted for by the proportion of exports in the total turnover (for more details, see Table 2 in the Appendix B). As recommended by Abadie *et al.* (2010) and Kaul *et al.* (2018), we add lagged outcome variables. For each outcome variable, we added six years of lagged outcome variable (2006, 2007, 2008, 2009, 2010, 2011).[17] Table 3 in appendix B displays the average values for the treated and control group for these ten selection variables.

Based on these ten predictive variables and lagged outcome variables, we built the synthetic control of each treated company as a convex combination of the 22 large companies in the donor pool. Due to confidentiality issues, we cannot present comparative statistics on the pre-treatment characteristics of each treated company, its synthetic version and those of all 22 large companies in the donor pool. In summary, we found high quality adjustments between two out of four companies and their synthetic controls. The adjustment of one of the remaining companies was less precise, while for the final company, the adjustment was of very poor quality, in that its synthetic control provided pre-treatment values that were very far from the ones this company had prior the implementation of the TRI programme. This suggests that the synthetic control method could not generate a suitable synthetic unit for this company. This may be an illustration of the "interpolation bias" as explained in *subsection III.2*. Therefore, the remainder of this analysis focuses on three firms only. For confidentiality reasons regarding their individual features, we have labelled them A, B and C.

Figure 2 (and respectively Figures 3 and 4) shows the evolution of R&D indicators for Company A (and respectively Companies B and C) and its synthetic counterpart over the 1998–2015 period. We can see that for each of these indicators, the trajectory of synthetic company A (and respectively of synthetic companies B and C) very closely follows the trajectory of Company A (and respectively Companies B

[17] In Abadie et al. (2010), the authors added three years of lagged outcome variables. No theoretical or empirical justification was put forward. However, as stated by Kaul *et al.* (2018) all the pre-intervention outcomes should not be used.

and C) over the whole period prior to implementation of the TRIs. This suggests that synthetic company A (and respectively synthetic companies B and C) provides a reasonable approximation of the evolution of the R&D indicators that Company A (and respectively Companies B and C) would have faced in the absence of the creation of the TRIs, although it can still be noted that the adjustment of C is worse than that of A and B. Turning to the post-treatment period then inform us on the impact of the policy.

Figure 2: Trends in R&D input indicators: Company A compared to synthetic company A

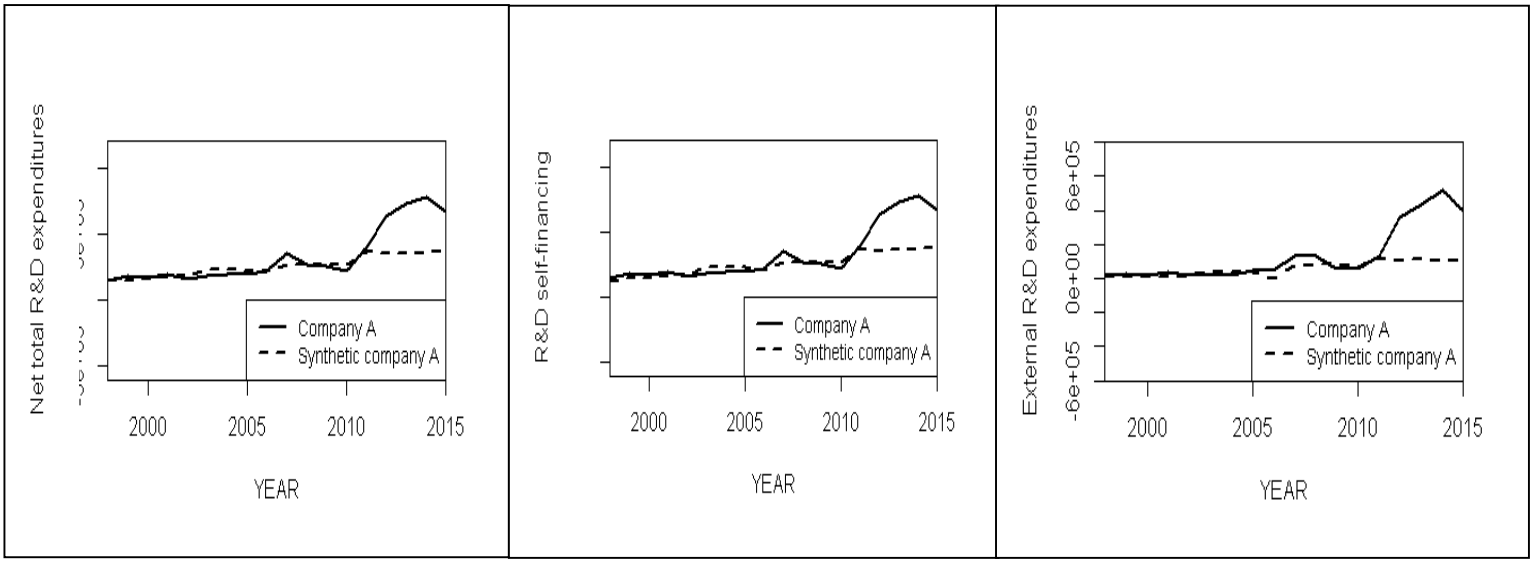


Figure 3: Trends in R&D input indicators: Company B compared to synthetic company B

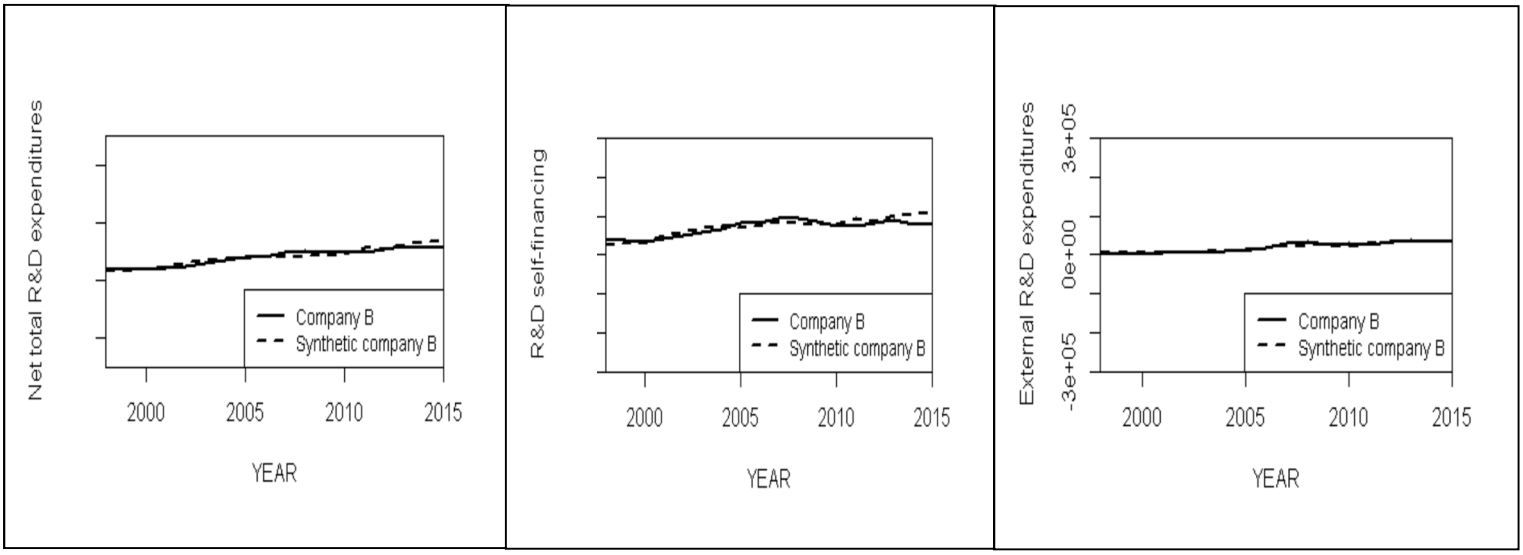
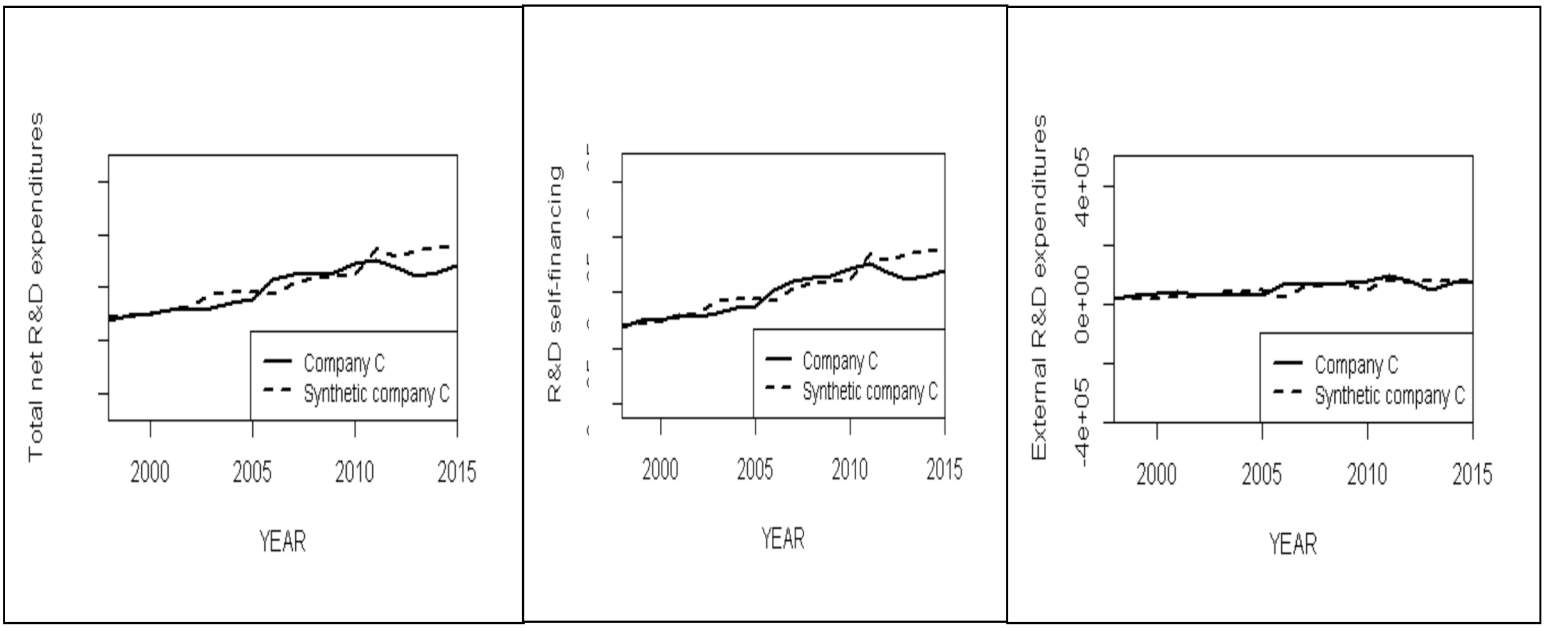


Figure 4: Trends in the R&D input indicators: Company C compared to synthetic company C



V.2. Individual impact estimation of the TRI programme

The impact of the TRI programme on the R&D indicators of a company is estimated by the difference between the values of these indicators of the company and the values of these same indicators of the synthetic control of the company, after implementation of the TRIs.

Regarding Company A, Figure 2 shows that for each R&D indicator, the two trajectories begin to considerably diverge as early as 2011 (one year before the creation of TRIs). While the trends in the R&D input indicators of synthetic Company A continued to increase slightly, those of Company A drastically increased. The divergence of the two trajectories one year before the effective creation of the TRIs may be interpreted in several ways. It could suggest the presence of the anticipatory effects of company A, which could be justified by the fact that the launch of the activities of these TRIs was largely the work of large companies, which had to invest, for example, in the acquisition of machinery and equipment, prior to the actual launch of the TRI's activities. However, it is also possible that the company simply changed its strategy by reinforcing its R&D capacity and adopting an "open innovation" strategy and that this change coincided with the year of preparation for entry into the TRIs. The estimations provided in Table 7 and the inference made in the next subsection will shed more light on this effect.

Conversely, for Company C, after treatment, performance in terms of R&D investment seems to be weaker than what could be expected based on the synthetic control performances. This is particularly marked for self-financed R&D, suggesting a windfall effect for this firm (see Figure 4).

The situation is more contrasted with regard to Company B. Analysis of Figure 3 shows that for external R&D expenditure, the two trajectories do not diverge significantly after the launch of the TRIs in 2012. The absence of a gap between the two trajectories suggests that the participation of Company B in the

TRI does not seem to have influenced the cooperative behaviour of this firm. In contrast, however, a negative effect is observed for self-financing indicators, which also indicates a potential windfall effect for this firm.

As recorded in table 7, the estimation of the annual average effect of the TRI on Company A shows that for the entire 2012–2015 period, net total R&D expenditure, external R&D expenditure and R&D self-financing are higher on average by almost 94%, 296% and 99% respectively, relative to the synthetic control (see Table 7). Regarding Company B, external R&D expenditure exceed on average by 7% those of the synthetic control. Net total R&D expenditure and R&D self-financing were reduced by the effect of the TRI on average by 8% and 18% respectively, relative to the synthetic control (see Table 8). With regard to Company C, net total R&D expenditure, R&D external expenditure and R&D self-financing are smaller than the one observed for the synthetic control by an average of nearly 22%, 14% and 21% respectively (see Table 9). However, it can be asked to what extent these estimated effects are credible. In other words, are they significant? This issue is addressed in the next paragraph.

Table 7: Annual gaps in R&D input indicators between Company A and its synthetic version

Post-treatment year	Gaps in total net R&D expenditure[1]	Gaps in R&D self-financing[1]	Gaps in external R&D expenditure[1]
2012	79.6%	79.4%	253.3%
2013	104.5%	99.7%	295.7%
2014	112.5%	110.6%	379.7%
2015	78.1%	74.8%	256.5%
Annual average effect	93.7%	99.1%	296.3%

1: Gaps in % are obtained by computing the ratio of the difference between treated firm's R&D input indicator and that of its synthetic version, and the synthetic version's R&D input indicator.

Table 8: Annual gaps in R&D input indicators between Company B and its synthetic version

Post-treatment year	Gaps in total net R&D expenditure[1]	Gaps in R&D self-financing[1]	Gaps in external R&D expenditure[1]
2012	-1.9%	-6.8%	7.1%
2013	-4.3%	-11.7%	17.8%
2014	-13.1%	-24.4%	-7.2%
2015	-12.3%	-31.2%	12.3%
Annual average effect	-7.9%	-18.5%	7.5%

1: Gaps in % are obtained by computing the ratio of the difference between treated firm's R&D input indicator and that of its synthetic version, and the synthetic version's R&D input indicator.

Table 9: Annual gaps of R&D input indicators between company C and its synthetic version

Post-treatment year	Gaps in total net R&D expenditure[1]	Gaps in R&D self-financing [1]	Gaps in external R&D expenditure[1]
2012	-13. 8%	-12.9%	1.0%
2013	-27.5%	-26.7%	-39.3%
2014	-25.8%	-25.1%	-8.5%
2015	-21.1	-20.5%	-9.3%
Annual average effect	-22.1%	-21.3%	-14.0%

1: Gaps in % are obtained by computing the ratio of the difference between treated firm's R&D input indicator and that of its synthetic version, and the synthetic version's R&D input indicator.

V.3. Inferences about the effects of the TRI programme

To assess the significance of these estimates within the context of the SCM, the research question asked was whether the results obtained could be entirely attributable to chance. In other words, how many times could the same results be obtained if a control company was randomly selected for the study in place of the treated company? To answer this question, we carried out a permutation test as described in *subsection III.3.*

Figures 5, 6, and 7 show the results of the permutation test associated respectively with Companies A, B and C. The green lines represent the difference associated with each of the 22 test runs (corresponding to the 22 control companies). In other words, the green lines show the difference in R&D indicators between each company in the donor pool and its respective synthetic version. The black line indicates the estimated gap for the treated company. With regard to Company A, one can see that for all R&D indicators, the estimated gap for the treated company during the 2012–2015 period is unusually large compared to the distribution of gaps for companies in the donor pool. The TRI programme appears therefore particularly effective for this firm both in terms of R&D investment and of cooperative behaviour. In contrast, there is no such a deviation in the trajectories of R&D indicators for Company B, suggesting that the TRI programme has had no significant additional nor windfall effect on input indicators for this firm. For Company C, on the other hand, we observe a clear negative impact of the TRI programme for all R&D input indicators except external R&D expenditure. This suggests that Company C's participation in the TRIs resulted in a windfall strategy substituting public finance to its own private investment in R&D. This may be due to better access to the research carried out by public labs involved in the TRIs, which reduces the need for internal R&D. Explanation which could be confirmed by the fact that company C's capacity to cooperate through the financing of external R&D isn't impacted by the TRI programme.

Our results, therefore, point to a strong heterogeneity in policy impact, underlining one of the main interests of the SCM that allows individual impact evaluation.

Figure 5: Gaps in R&D input indicators and placebos in all 22 control companies

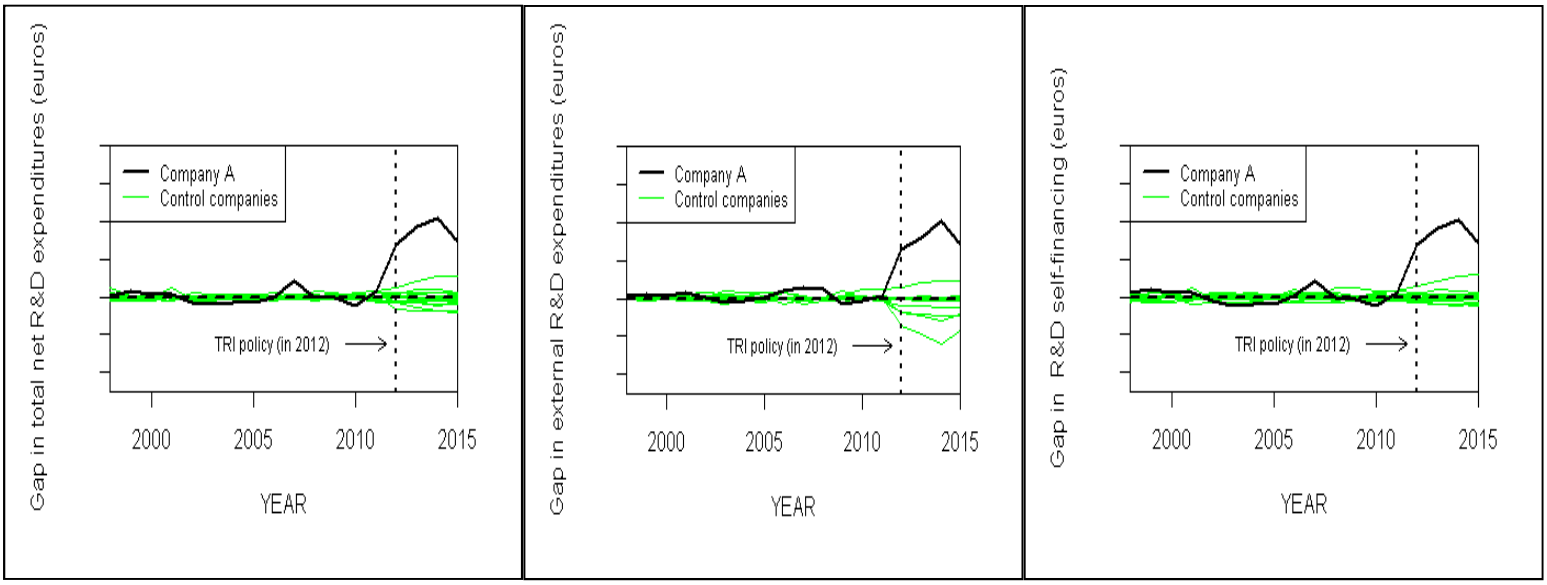


Figure 6: Gaps in R&D input indicators and 22 placebo in all 22 control companies

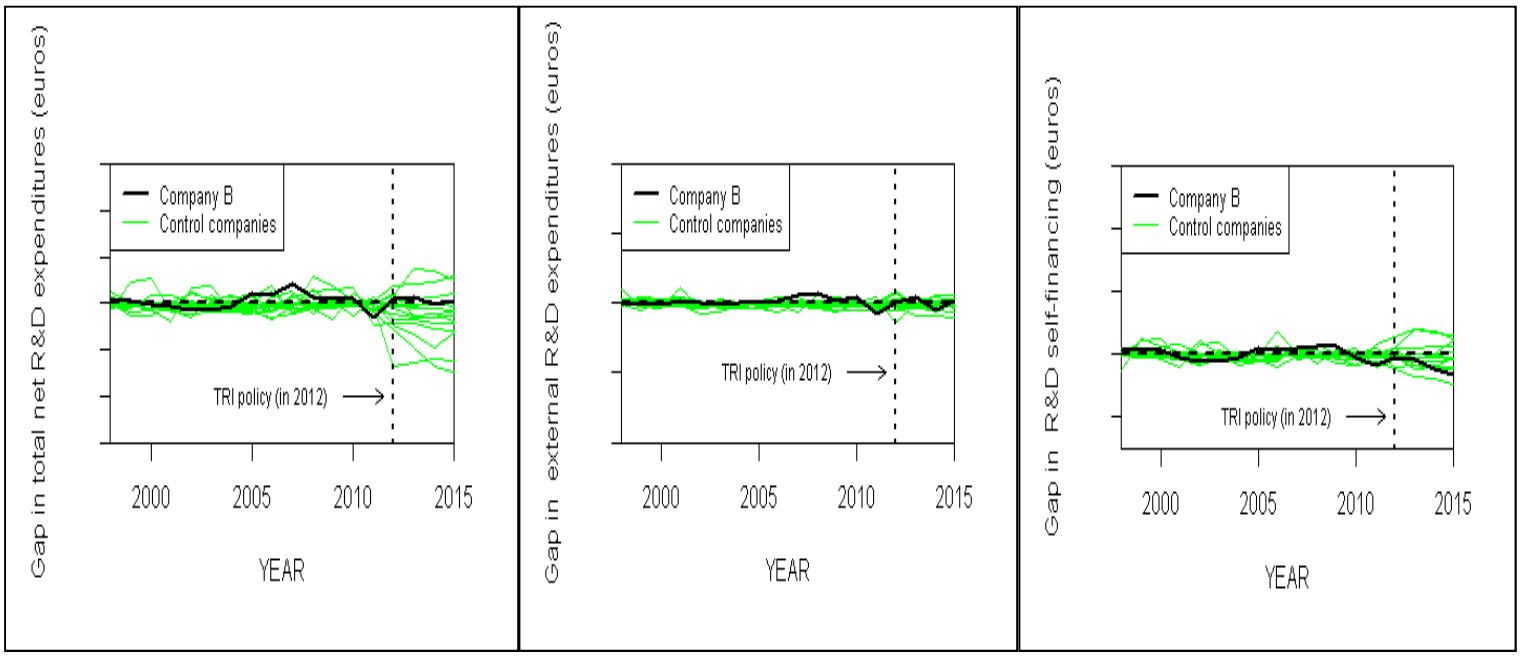
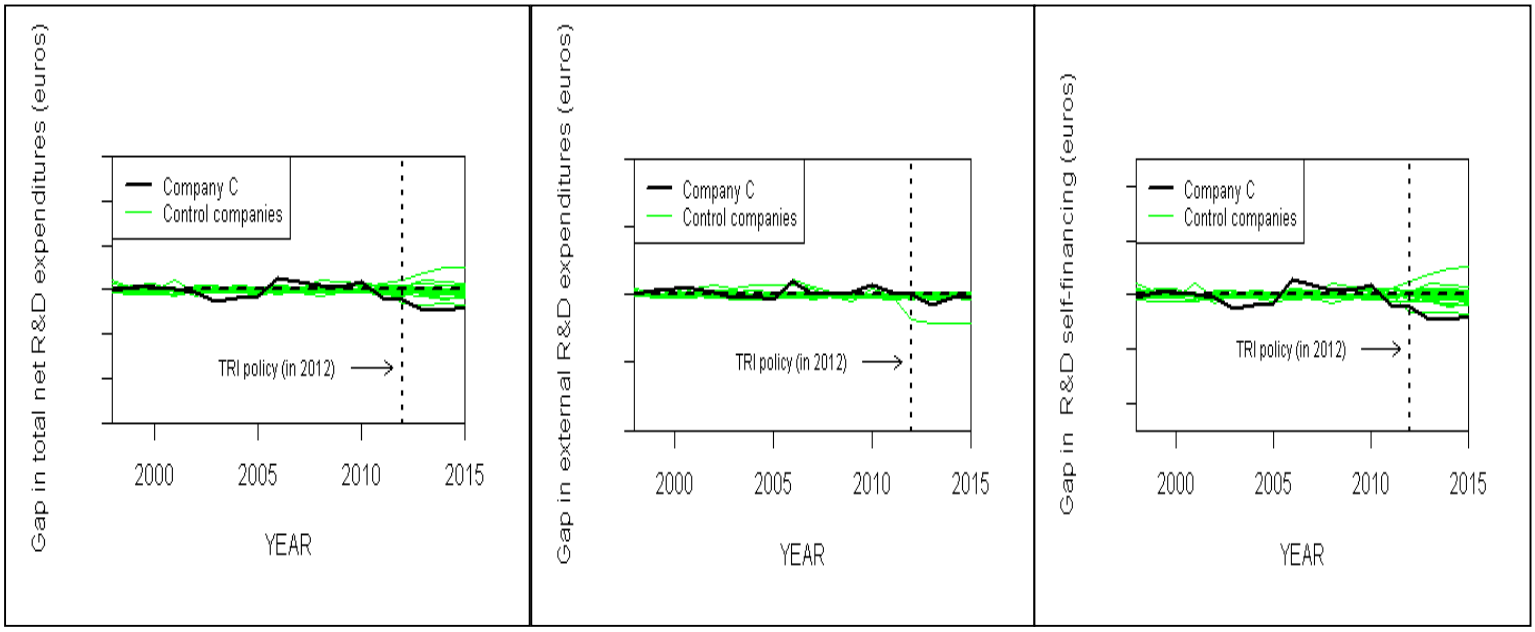


Figure 7: Gap in R&D input indicators and placebos in all 22 control companies



V.4. Comparison between the synthetic control method and traditional evaluation methods

In order to better grasp the contribution of the SCM for evaluating the impact of public programmes on large firms, we compare the results obtained by the SCM with those obtained when applying the most standard impact evaluation approach used within the existing literature, the Difference-in-Differences method (DiD). Taking advantage of the panel structure of our data, we also estimate impact of fixed effect (FE) models to control for individual heterogeneity. In addition, although not yet used so for microeconomic analysis, we apply the random trend model (Wooldridge, 2002) to our case. Indeed, as it enlarges the capacity of the model to control for individual effects by introducing an individual trend, this model seems also particularly interesting for evaluating the impact on large firms, which generally follow specific trajectories. It allows us to go beyond individual fixed effects and temporal fixed effects, to take into account a second source of heterogeneity, the linear individual trend, which refers to firm's temporal events (mergers, acquisitions, disposals, etc.) which can influence the current performance of each large company, even in the absence of public policies. Moreover, large companies, unlike some small companies, invest permanently in R&D, whether or not they benefit from innovation policies. Therefore, the estimated effects may be tainted by a bias due to the presence of a tendency to invest in R&D activities.

It is worth noting, however, that these three methods aim at measuring the average effect of the TRI programme, whereas the SCM allows us to estimate individual effects, i.e., effects measured for each

firm taken separately. To make comparison possible, we also calculated an average treatment effect[18] summarising the result obtained with the SCM.

Table 10 presents a summary of the results obtained using the four different methods. In order to compare the SCM to the other methods, we computed the average treatment effect of the three companies A, B and C. Note that it is not possible to achieve statistical significance with this average SCM effect, contrary to what was achieved with the other methods.

Table 10: Method comparison: estimated effects using the difference-in-differences method, the fixed effects model, the random trend model and the SCM

Outcome variables	D-i-D	FE model	Random trend model	SCM
Net total R&D expenditure	313,716 (233,544)	106,768*** (38,910)	80,376*** (18,494)	84,341
External R&D expenditure	415,798 (850,232)	87,042*** (36,008)	69,253*** (10,921)	103,879
R&D self-financing	270,714 (211,924)	100,145** (40,093)	80,197*** (18,272)	80,570

*Significance at *** 1% threshold; ** 5% threshold and * 10% threshold. Standard errors in parentheses.*

The results obtained using the DiD method seem to indicate a non-significant effect of the TRIs on all the outcome variables. This is probably due to the strong inter-individual heterogeneity of large firms. The results of the FE model show smaller but significant estimated impacts for the three outcome indicators. Including time and individual fixed effects, this method controls for the specific firm features or temporal shocks that explain both participation in the TRI and the R&D performances, allowing a more accurate assessment of policy impact.

The random trend model, generally used in regional impact assessments such as the SCM, goes a step further, making it possible to take into account, beyond individual fixed effects and fixed temporal effects, a second source of heterogeneity, which is the linear individual trend specific to each company. This second heterogeneous source is very common among large companies insofar as they have strategic behaviours (mergers, acquisitions, disposals, etc.) that change over time, which can influence the current performance of large companies, even in the absence of public policies. Moreover, large companies, unlike some small companies, invest permanently in R&D, whether or not they benefit from innovation policies. Therefore, the estimated effects may be biased due to the presence of a tendency to invest in R&D activities.

By extracting the individual trend specific to each large company through application of the random trend model, significant effects are still observed on the three R&D indicators, albeit with a smaller magnitude. The effect on total net R&D expenditure is about €80 million; the effect on external R&D expenditure is estimated at about €69 million, while for self-financing, the treatment effect is about €80 million. The increase in total R&D therefore seems to be mainly self-financed and subcontracted.

[18] The average treatment effect measures the difference in mean (average) outcomes between units assigned to the treatment and units assigned to the control.

Interestingly, these random trend model effects are not very different from those obtained by averaging the estimated effects from the SCM. In contrast, the DiD method and the FE model tend to overestimate the effects of the policy, compared to the SCM and random trend model. Compared to the FE and random trend models, the SCM concludes with a more pronounced impact on external R&D expenditure. This could indicate that giving the same weight to all companies in the donor pool in the control sample does not allow for precise measurement of the impact of the policy on large firms participating in TRIs.

Moreover, as detailed in the previous paragraph, the SCM has the advantage of allowing a detailed analysis of the heterogeneity of the impacts. Contrary to the FE and random trend models, it reveals that not all companies benefit from the TRI programme in the same way. In our case, the positive average treatment effect is only due to one company, Company A, showing systematic significant effects on all of its R&D input indicators. For the other two companies, no significant change in R&D behaviour is identified and some windfall effects may even appear.

VI. Conclusion

Large companies generally receive a large proportion of public aid compared to other categories of companies, particularly SMEs. Identifying the quantitative impact of public policies on large companies is therefore crucial for public authorities and a challenge for researchers. However, due to the singularity of these companies and the small sample size, it is difficult, if not impossible, to find relevant counterfactual units for large companies in order to conduct a rigorous impact assessment. The main contribution of this article is, therefore, to propose and explore the relevance and feasibility of implementing a different evaluation method, namely the synthetic control method, to assess the impact of public policies on the performance of large companies. Initially designed to assess the impact of a major event or policy on a single aggregated treated unit (a city, region or country), the interest of this method lies in its ability to construct a valid counterfactual for a single treated unit and thus to estimate the individual effect of the treatment.

Based on firms' participation in a new science-industry transfer policy, the French Technological Research Institutes, we analysed four large companies' R&D investments before and after their participation in the policy. We found that the SCM succeeded in generating a reasonable counterfactual for three of the four companies involved in the TRIs under scrutiny. We estimate the difference-in-differences model and the mixed fixed-effects model to compare their results with those obtained with the SCM. In addition, we specify and estimate the random trend model that is also generally used in regional impact assessments.

In the end, it appears that the degree of convergence between evaluation methods depends on the quality of the control they perform. While no significant effects are found with the DiD method, the fixed effect model tends to overestimate the treatment effects compare to the random trend model. This is probably explained by the fact that these traditional evaluation methods are not able to take into account the weight and the own strategies, which have a very high impact on the performance of large companies.

Closer to the random trend model, the SCM gives very similar average impacts for total R&D and self-financed R&D but the assignment of different weights to the donor pool control sample leads to higher average estimated impacts on external R&D. This may be due to the significant heterogeneity of the treated and control companies over the pre-treatment period. Unobserved confounding factors may prevent the proper estimation of the impact of the policy on firms' R&D subcontracting when the random trend model is used. In addition, the SCM leads to a much better interpretation of the results by allowing assessing the policy impact at the individual level. Indeed, the average treatment effect hides great differences in behaviour between companies. A positive impact was found on the R&D inputs of one firm, but a negative impact was observed for a second firm and no significant effect appeared for the third. The significant impact on self-financing indicates a leverage effect for Company A and a windfall effect for Company C. Similarly, the TRI programme was only able to significantly influence the R&D cooperative behaviour measured by external R&D for Company A. The effect for Company B is also positive but not significant. The effectiveness of this new platform-based science-industry initiative in France, is therefore not yet clear. Its effect on large firms is highly heterogeneous, and one of the main reasons for using the SCM is to highlight this heterogeneity. Although for reasons of confidentiality we cannot develop this point further, our results allow the development of a more qualitative interpretation of the impact evaluation by looking for the reasons for these differences in the strategy of the participating firms. The modalities of each TRI implementation might also be crucial, since the design of the various platforms and the level of large firms' involvement on these platforms varies strongly. Thus, the differences in results between firms could be explained by the fundamental differences between the two TRIs (Nanoelec and Bioaster), particularly in terms of operating methods, research themes, etc. For example, in Nanoelec the private founding members (large companies) do not have to pay in order to carry out one or more R&D projects, whereas in Bioaster they are obliged to pay for the R&D service as if they were a simple client coming from outside the TRI.

The SCM is not without its limits. Often, implementation of the SCM faces the main problem of over-adjustment. This occurs when the characteristics of the unit affected by the intervention are artificially matched by combining idiosyncratic variations in a large sample of unaffected units (Abadie *et al.*, 2015). Moreover, our study point to another limit of the SCM. In some cases, the SCM may not be able to construct a suitable synthetic control for a treated unit and therefore it would be impossible to estimate an accurate impact of a policy because accuracy of an estimation depends on the quality of adjustment of the synthetic control.

In spite of these limitations, this study raises several potential avenues for research. Firstly, it would be interesting to analyse whether the observed increase in R&D investment and collaborative R&D has been transformed into innovative performance (patent applications, sales of innovative products, etc.) and, consequently, into socio-economic performance (turnover, employment, equity, etc.). Such an analysis will only be possible with a fairly significant time lag and provided that we are able to segment the effect of different public support for each large company. Furthermore, it would be interesting to more thoroughly analyse the behavioural additionality of large firms with respect to their participation in the TRIs by using new indicators. This would involve examining whether the TRI has influenced the R&D collaborative strategy of firms, especially between large firms and SMEs. More particularly, it could be investigated whether the participation in the TRI has improved their ability to collaborate with other firms and research centres in and outside the TRI and whether treated firms have launched riskier, high-potential R&D projects (OECD, 2006). Secondly, this study focuses only on the Rhone-Alpes TRIs, which means that the results cannot be generalised. Therefore, it would be important to evaluate the six other TRIs established in France in order to have a comprehensive view of the effectiveness of the TRI policy and to be able to develop interpretation of the heterogeneous effectiveness of the different ways that have been chosen to implement the TRIs.

Acknowledgment:

References

Abadie, A., Gardeazabal, J., 2003. The economic costs of conflict: A case study of the Basque country. American Economic Review 93(1), 112–32.

Abadie, A., Diamond, A., Hainmueller, J., 2010. Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. Journal of the American Statistical Association 105(490), 493–505.

Abadie, A., Diamond, A., Hainmueller, J., 2015. Comparative politics and the synthetic control method. American Journal of Political Science 59(2), 495–510.

Afcha, S., García-Quevedo, J., 2016. The impact of R&D subsidies on R&D employment composition. Industrial and Corporate Change, Oxford University Press 25(6), 955-975.

Angrist, J., Krueger, A., 1999. Empirical strategies in labor economics. Handbook of labor economics, 3 :1277–1366.

Angrist, J.D., Pischke, J.S., 2008. Mostly harmless econometrics: An empiricist's companion. Princeton university press.

Athey, S., Imbens, G. W., 2017. The State of Applied Econometrics: Causality and Policy Evaluation. Journal of Economic Perspectives 31(2), 3-32.

Auld, M. C., Grootendorst, P., 2004. An empirical analysis of milk addiction. Journal of Health Economics 23(6), 1117–1133.

Ben Hassine, H., Mathieu, C., 2017. Évaluation de la politique des pôles de compétitivité: la fin d'une malédiction? Document de travail France Stratégie.

BEIS, 2017. The impact of public support for innovation on firm outcomes. Department for Business, Energy and Industrial Strategy.

Boeing, Philipp., 2016. The allocation and effectiveness of China's R&D subsidies - Evidence from listed firms. Research Policy 45(9), 1774-1789.

Bondonio, D., Martini, A., Biagi, F., 2015. Counterfactual impact evaluation of enterprise support programmes. evidence from a decade of subsidies to Italian firm. 55th Congress of the European Regional Science Association: "World Renaissance: Changing roles for people and places", 25-28 August 2015, Lisbon, Portugal.

Clarysse, B., Wright, M., Mustar, P., 2009. Behavioural additionality of R&D subsidies: A learning perspective. *Research Policy* 38(10), 1517-1533. https://doi.org/10.1016/j.respol.2009.09.003

Costa Dias, M., 2014. The synthetic control method compared to difference in differences: discussion. in: Towards a common language in statistical methodology: examples from current applications in economics and health research.

Criscuolo, C., Martin, R., Overman, H. G., Reenen, J. V., 2012. The causal effects of an industrial policy. NBER Working Paper No. 17842.

DiNardo, J., Pischke, J., 1997. The returns to computer use revisited: Have pencils changed the wage structure too? The Quarterly Journal of Economics 112(1), 291–303.

Dimos, C., Pugh, G., 2016. The effectiveness of R&D subsidies: A meta-regression analysis of the evaluation literature. Research Policy 45(4), 797-815.

Dujardin, C., Louis, V., Mayneris, F., 2015. Les pôles de compétitivité wallons quel impact sur les performances économiques des entreprises? the walloon competitiveness clusters and their impact on firms' economic performances? IRES Discussion Papers.

Greco, M., Grimaldi, M., Cricelli, L., 2017. Hitting the nail on the head: exploring the relationship between public subsidies and open innovation efficiency. Technological Forecasting and Social Change 118, 213–225.

Heckman, J., Hotz, V. J., 1989. Choosing among alternative nonexperimental methods for estimating the impact of social programs: The case of manpower training. Journal of the American Statistical Association 84(408), 862–74.

Imbens, G.W., Rubin, D. B., 2015. Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction. G New York: Cambridge University Press.

Jugend, D., Fiorini, P. D. C., Armellini, F., Ferrari, A. G. 2020. Public support for innovation: A systematic review of the literature and implications for open innovation. Technological Forecasting and Social Change, 156.

Kaul, A., Klößner, S., Pfeifer, G., Schieler, M., 2018. Comparative politics and the synthetic control method revisited: a note on Abadie *et al.* (2015). Swiss Journal Economics Statistics 154(11).

OECD, 2006. Government r&d funding and company behaviour, measuring behavioural additionality. OCDE.

OCDE, 2015. OECD Science, Technology and Industry Scoreboard 2015: Innovation for growth and society, Éditions OCDE, Paris, https://doi.org/10.1787/sti_scoreboard-2015-en/

Saia, A., 2017. Choosing the open sea: The cost to the UK of staying out of the euro. Journal of International Economics 108(C), 82-98.

Stuart, E. A., 2010. Matching methods for causal inference: A review and a look forward. Statistical science: a review journal of the Institute of Mathematical Statistics 25(1), 1-21.

Vanino, E., Roper, S., Becker, B., 2019. Knowledge to money: assessing the business performance effects of publicly-funded R&D grants. Research Policy 48, 1714-1737.

Wanzenböck, I., Scherngell, T., Fischer, M. M., 2013. How do firm characteristics affect behavioural additionalities of public R&D subsidies? Evidence for the Austrian transport sector. *Technovation* 33(2-3), 66-77. https://doi.org/10.1016/J.TECHNOVATION.2012.11.006 [Links]

Wooldridge, J., 2002. Econometric Analysis of Cross Section and Panel Data, MIT Press.

Appendix A: Comparative analysis between the treated firms and the donor pool sample

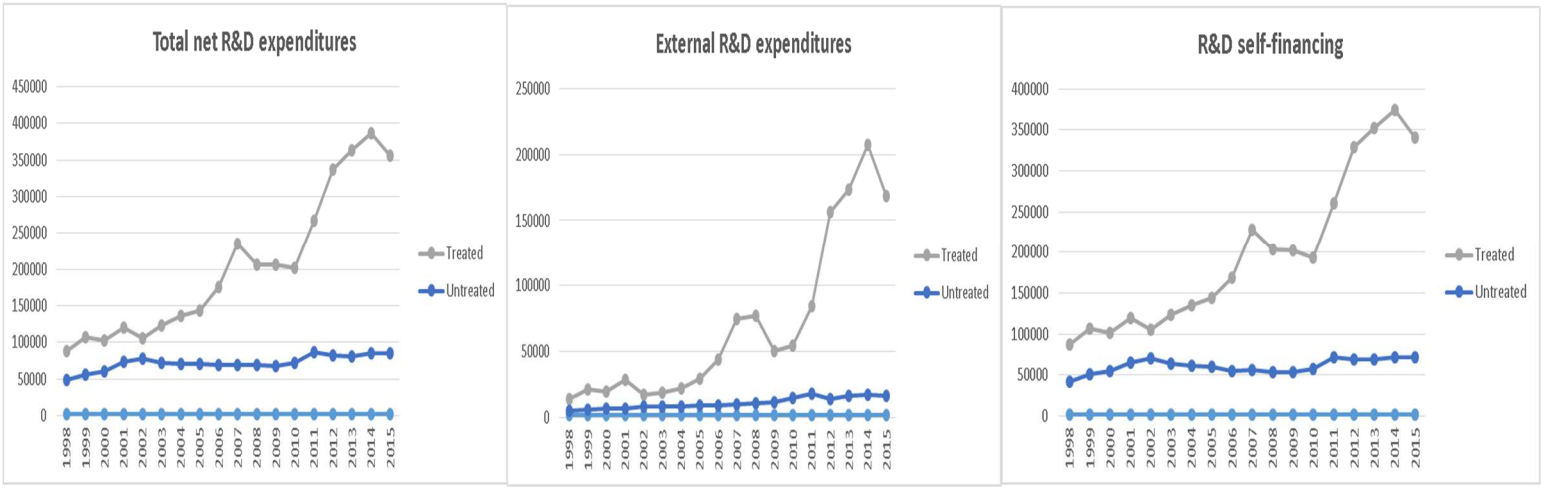
Table 1: Outcome comparison between treated and control companies over the pre-treatment period (1998–2011)

Outcome variables	Treated group	Control group	Difference
Net total R&D expenditure[1]	232,476	68,815	163,661***
R&D self-financing[1]	229,819	58,517	171,302***
External R&D expenditure[1]	42,590	9,792	32,798***
Number of companies	4	22	

*1.In thousands of euros. Significance at *** 1% threshold; ** 5% threshold and * 10% threshold.*

Table 1 clearly shows that the treated and control groups are not statistically comparable before the creation of the TRIs, as the differences between these two groups in terms of outcome are statistically significant.

Figure 1: Annual average evolution of R&D input indicators (treated vs untreated)



As Figure 1 shows, the treated and control groups do not have the same pre-treatment trajectories, regardless of the indicator. In addition, we note the presence of peaks and troughs which are often common to the trajectories of both treated and untreated firms, suggesting the existence of fixed temporal effects and sometimes specific to the trajectories of treated firms, which may suggest the existence of individual annual shocks. These peaks and troughs could also reflect the effects of the strategies (mergers, acquisitions, disposals, etc.).

Appendix B: Variable description

Table 2: List of dependent and independent variables

	Description	Unit
Dependent Variables		
Total R&D expenditure net of public funds	Total R&D expenditure minus public funds.	In thousands of euros
External R&D expenditure	All R&D expenditure outside the company in the form of collaboration and subcontracting.	In thousands of euros
R&D self-financing	Financing from the company's own funds	In thousands of euros
Predictive Variables		
Total turnover	Total amount of turnover at the firm level	In thousands of euros
Equity	Sums paid by the partners or shareholders, plus the profits generated annually by the company which are not distributed as dividends.	In thousands of euros
Total number of employees	Total number of employees.	In terms of individuals (not full-time jobs)
Public financing	Total amount of R&D funding from the public sector (subsidies, etc.).	In thousands of euros
Private financing	Total amount of R&D funding from the private sector (private companies, etc.).	In thousands of euros
Capital R&D expenditures	Total amount of R&D expenditure related to the acquisition of R&D machinery and equipment.	In thousands of euros
Research credit tax	Total amount of R&D tax credit the firm received	In thousands of euros
Proportion of high skilled labour force	Ratio between the number of managers and the total number of employees.	
Proportion of exports	Ratio between exports and turnover.	
Number of establishments	The number of establishments owned by the large company.	
Treatment Variable		
TRI participation	Dummy variable which takes the value 1 in 2012, 2013, 2014 and 2015 if the company has been involved in one of the two TRIs in the Rhône-Alpes and 0 otherwise	

Table 3: Comparison of predictors between treated and control companies over the pre-treatment period (1998–2011)

Predictive variables	Treated group	Control group	Difference
Turnover[1]	1,591,081	1,326,125	264,956
Number of employees [2]	6,419	5,261	1,158*
Equity[1]	1,540,272	781,748	758,524***
Public funds[1]	15,582	299	15,283***
Private funds[1]	2,657	10,044	-7,387***
Capital R&D expenditure[1]	39,079	7,571	31,508***
Proportion of executives	0.289	0.272	0.017
Proportion of exports	0.711	0.566	0.144***
Number of establishments	18	19	-1
Number of companies	4	22	

*1: In thousands of euros; 2: In number of individuals. Significance at *** 1% threshold; ** 5% threshold and * 10% threshold.*